《心理学报》论文自检报告

请作者填写以下内容, 粘贴在稿件的首页。

1. 请以“研究亮点”的形式列出最多三条本研究的创新性贡献, 总共不超过 200 字。

《心理学报》的目标是发表“既科学优秀, 又具有广泛兴趣和意义” (be both scientifically excellent and of particularly broad interest and significance)的心理学前沿研究。如果您的研究只有小修小补的贡献, 没有尝试开创新的研究领域(new areas of inquiry)或提出独到见解和创新视角(unique and innovative perspectives), 特别纯粹只是研究没有明确心理学问题的算法或技术的工作, 这类研究被本刊接受的机会小, 建议另投他刊。

答：（1）首次发现概率类数字信息框架会降低人们对 AI 的威胁感，并且增加人们对 AI 研发政策的支持以及增加对高 AI 曝光度的行业工作岗位的推荐度，揭示了“AI 失业”背景下的信息框架效应；（2）发现主观可能性判断在概率信息框架降低 AI 威胁感中起中介作用，揭示了“AI 失业”信息框架效应的潜在机制；（3）发现个体模糊容忍度的调节作用，有助于更好地理解与应对 AI 失业相关的负面社会结果。

2. 作者已经投稿或发表的文章中是否采用了与本研究相同的数据？如果是, 请把文章附上审查。(我们不赞成作者用同一数据发表多篇变量相同的文章, 也不赞成将一系列的相关研究拆成多个研究来发表的做法。)

答：否。

3. 管理、临床、人格和社会等领域仅有自我报告(问卷法)的非实验非干预研究, 需要检查数据是否存在共同方法偏差(common method bias)。为控制或证明这种偏差不会影响研究结论的效度, 你使用了什么方法？采取了哪些措施？(共同方法偏差的有关文献可参见：http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract894.shtml)基于横断数据, 仅有自我报告, 仅仅在方便样本中施测, 这样的研究数据易取得, 但通常创新性价值不大, 被本刊接受的机会小。

答：非纯问卷研究。

4. 是否报告并分析了效果量(effect sizes; 如：t 检验：Cohen’s d; 方差分析：η^2 或 η_p^2; 标准化回归系数)？(很多研究只是机械地报告了效果量, 但没有做必要的分析或说明, 如效果量是大中小？有什么理论意义或应用意义？)。(在 google 中搜索“effect size calculator”, 可搜到许多计算方便的 APP。效应量的有关解释, 中文可参考：http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract1150.shtml；英文可参看：http://www.uccs.edu/lbecker/effect-size.html

是否报告统计分析的 95% CI？(如, 差异的 95% CI; 相关/回归系数的 95% CI)置信区间的有关计算和绘图可参考 https://thenewstatistics.com/itns/esci/)

答：是。

5. 请写出计划的样本量, 实际的样本量。如果二者有差别, 请写出理由。以往心理学研究中普遍存在样本量不足导致的低统计功效(power)问题, 我们建议在论文的方法部分解释您计算及认定样本量的依据。应该以有一定依据的效果量(effect size)、期望的功效来确定样本量, 并报告计算用软件或程序。样本量计划的理由和做法可参考 https://osf.io/5awp4/

答：已计算，并写明。

6. 假设检验中, 如果是零假设显著性检验(NHST), 需报告精确 p 值而不是 p 的区间(小于 0.001 的报告区间, 其他报告精确 p 值)。你的论文是否符合该项要求？如果是贝叶斯因素, 是否已报告其对先验分布假定的敏感性？

答：符合要求，均已报告。

7. 为保证论文中数据报告的完备性, 统计分析中如果剔除了部分数据, 是否在文中报告？原因是什么？包含这部分数据的统计结果如何变化？统计分析中是如何处理缺失数据的？使用量表时是否删除了其中的个别题目？原因是什么？如果包含这部分题目, 统计结果会如何变化？是否有测量的项目或者变量没有报告？原因是什么？请写出在论文中的位置。

答：实验研究通过注意检查剔除被试，已报告。

8. 研究用到的未经过同行评议和审查的实验材料、量表或问卷, 是否附在文件的末尾以供审查？如果没有, 请写出理由。如果该文发表, 您是否愿意公开这些材料与其他研究者共享？

答：是，文中已附。

9. 本刊要求作者提供原始数据, 请在以下 3 种里选择一种打√:

a)投稿后将数据发至编辑部邮箱 (√)

b)数据可以从如下链接中获得 ________________________ ( )

c)原始数据和程序已在科学数据银行(https://www.scidb.cn/surl/xlxb)上分享 ( )

d)如不能提供，请说明理由或提供有关证明。

10. 您的研究是否是临床干预或实验室实验？ 是( ) 否(√)

如果是, 请提供预先备案登记号 ___________。

如果没有, 请说明原因____________________。

注：临床干预或实验室实验, 建议在收集数据前预先备案(pre-register)。也鼓励其他实验研究预先备案。备案要求写出所有的研究假设及其支持, 以及实验/干预的详细过程和步骤。本期刊的预先备案网站 https://osf.io/ 或 https://aspredicted.org/。如果您的研究有预先备案, 会显著增加被录用的机会。预先备案的重要性可参考 https://osf.io/5awp4/ 。

11. 您的研究如果用到了人类或动物被试, 是否得到所在单位伦理委员会的批准？如果是, 请把扫描版发至编辑部邮箱。如果否, 请说明理由。

答：是，好的。

12. 是否依据编辑部网站发布的“英文摘要写作注意事项”撰写 400~500 个单词的英文大摘要？英文题目和摘要是否已请英语好的专业人士把关或者已送专业 SCI/SSCI 论文编辑公司修改润色？

答：已按要求撰写并请人把关。

13. 如果第一作者是学生, 请导师单独给编辑部(xuebao@psych.ac.cn)发邮件, 说明已阅读本文并认真把关。是否已提醒导师给编辑部发邮件？(编辑部收到导师邮件后才会考虑进入稿件处理流程)

答：第一作者非学生。

14. 请到编辑部网站首页右侧“下载中心”下载并填写“稿件不涉密证明”, 加盖通讯作者单位的保密办公章, 把扫描件发至编辑部邮箱(xuebao@psych.ac.cn)。如没有保密办公章, 请加盖通讯作者的单位公章。是否已发邮件？

答：是。

“AI 失业”的信息框架效应

摘 要 人工智能技术的发展引起人们对未来失业问题的关注，而在了解“AI 失业”过程中所接收的信息则可能影响人们的 AI 威胁感。通过七个实验，文章探讨了关于“AI 失业”的两种数字信息框架对 AI 威胁感的影响和其内在机制和边界条件、以及对 AI 研发支持和工作评价标准等后续影响。结果发现：相比于因素框架，概率框架降低了人们的 AI 威胁（实验 1–7）；其机制为判断“AI 失业”发生可能性的中介作用，即概率框架会使人们形成“AI 失业”更不可能发生的判断，从而降低 AI 威胁（实验 2–5）；此效应受到模糊容忍度的调节，即概率框架降低 AI 威胁主要体现在对模糊信息容忍度高的个体当中（实验 5）；此外，受概率框架影响，人们也对有关 AI 研发的政策表现出更加支持的态度，AI 威胁在其中发挥中介作用（实验 6）；同时，受概率框架影响的人们对于需要经常接触 AI 的工作也有更加积极的推荐意愿（实验 7）。研究结果区分了两类数字信息框架对 AI 威胁的影响，为以往探索 AI 威胁成因的研究提供了新证据。

关键词 技术性失业；信息处理；概率判断；AI 威胁；模糊容忍度

1 引言

人工智能（Artificial Intelligence，以下简称 AI）被誉为“本时代最重要的通用技术”，（Felten et al., 2021），无论程序编写还是疾病诊断，无论生成图像还是日常聊天，AI 在诸多领域比拟人类的表现引来媒体的大量报道，人们也由此感受到其巨大的应用潜力（Metz, 2022; Roose, 2022; Savage, 2020; Strogatz, 2018; Verma, 2022）。据统计，2022 年以来，美国主流刊物平均每 200 篇报道中便有 1 篇以“AI”为主题，而 2012 年时，每 1150 篇中才会出现 1 篇关于 AI 的报道（Santoro & Monin, 2023）。然而，在大众不断接触 AI 技术的过程中，一些社会问题也随着高频曝光而愈演愈烈（Brynjolfsson & McAfee, 2017）。例如欧美国家食品、运输、销售等行业频繁出现的抵制自动化罢工，欧盟制定严格的准入规范以限制 AI 技术在生产端的投入，而近期 Duolingo（全球最大的外语学习社区）引入 AI 产品并裁撤千名翻译员工事件，以及职业画师发起的“拒绝 AI 绘画（No To AI Generated Images）”运动，则引起了民众对 AI 的广泛质疑和警惕（Hsu, 2022; Iza World of Labor, 2019; Nemo, 2018; Sparrow, 2007; Złotowski et al., 2017）。此类事件中，AI 和人的矛盾集中于工作职位中潜在的竞争关系，“AI 失业”成为舆论关注的焦点话题（Wike & Stokes, 2018）。

AI 失业属于技术性失业的范畴，即 AI 技术在对人力劳动形成替代效应、冲击传统就业结构、夺走就业机会后造成的大规模失业（Schlogl & Sumner, 2018）。由于目前 AI 技术

尚处于由“弱 AI”至“强 AI”的研发阶段，一般认为“AI 失业”这一概念是对未来就业环境的预测（许丽颖, 喻丰, 2020; Makridakis, 2017）。而在 AI 失业是否及以何种规模发生这一问题上，却一直存在截然不同的两类观点：基于市场调节和补偿机制的有关理论，持乐观态度者认为，AI 与以往的变革性技术一样，在淘汰旧有岗位的同时也能创造出更多新岗位（Makridakis, 2017; Padalino & Vivarelli, 1997）；但考虑到 AI 技术的发展目标是开发出模拟人类智能的机器智能，也有人预测 AI 造成的替代效应要多于创造效应（Frey & Rahbari, 2016）。

与理论预测的分歧相似，针对 AI 失业的调查也未能得出一致的结论。麦肯锡全球研究的测算结果显示，美国 51%的有薪岗位将在 20 年内被 AI 逐渐取代（Chui et al., 2016），这一数字在中国为 77%（范长煜, 邓韵雪，2022）；但关于 AI 替代效应的调查结果显示，已发生的“机器换人”实际比率在美国仅为 0.18%–0.34%（Acemoglu & Restrepo, 2020），在中国约为 2.5%（程虹 等, 2018）。另有研究发现，西班牙制造业在引入 AI 后总计创造出了 10%的新岗位（Koch et al., 2021），但在非制造业使用 AI 却又造成了就业人口减少 0.2%且工资水平下降 0.37%的结果（Acemoglu & Restrepo, 2020）。

以上分歧反映出人们对“AI 失业”的认识中存在的不确定性，也意味着人们对 AI 失业乃至 AI 本身的态度均可能被相关信息所引导（Blanas et al., 2019; Dauth et al., 2018）。对多数人而言，传播媒介对 AI 失业所涉行业和 AI 失业规模概率的报道是其了解此事的主要途径（Morikawa, 2017）。而该过程中可能存在信息框架效应，即对同一事件的不同表述、有侧重地凸显其部分特点，会引起人们不同的反应（Levin et al., 1998）。因此，即便是关于 AI 失业的相同内容，即便采用的是同一批客观数据，置于特定的信息框架也可能引起民众截然不同的反应，尤其是“AI 威胁（perceived AI threat）”，即因 AI 而生的不适、焦虑和恐惧感（Gray et al., 2024; McFarland, 2014; Złotowski et al., 2017）。

与一般性的威胁感知一样，AI 威胁感只反映人类对 AI 的心理知觉结果，并非以充分的客观事实为基础（Gray et al., 2024），这也解释了即便目前 AI 对就业的实际影响尚不明确，AI 威胁感便已普遍存在于不同的文化与地区中（Lingmont & Alexiou, 2020; Yam et al., 2023）。以 2015 年的调查结果为例，美国有近 30%的受访者担心被 AI 取代（McClure, 2018）；而在 2019 年国内调查的结果中，约 46%的受访者表达了不同程度的担忧（李佩, 2019）。虽然 AI 威胁的形成不是全然理性的结果，但又对人们以何种方式、程度运用 AI 技术产生影响，同时也影响个人职业规划方面的决策及行为（Dekker et al., 2017; McClure, 2018）。

本文所探讨的问题在于，人们在了解 AI 失业的过程是否受到信息框架的影响？信息框架效应是否表现为 AI 威胁感知的差别？在此基础上，人们对 AI 技术的态度以及个人就业的决策行为又会发生怎样的变化？此种信息框架效应是否也受到特定个体特征的影响？为此，本文通过 7 个主要实验考察以“AI 失业”为对象的信息框架对 AI 威胁感知的影响和

机制，并为 AI 在信息传播等方面的治理问题提供一定的实践启示。

1.1 基于因素和概率的信息框架效应

人们通过事物传递的信息完成决策和行为（Betsch et al., 2015; Brewer et al., 2004; Diener & Richardson, 2007）。然而，人类的认知加工方式决定其只能获取事物的部分信息，导致对同一对象的理解可能存在差异，这种可能性成为信息框架效应存在的基础（Anderson, 1965）。信息框架效应（message framing effect）指不同的信息表述方式导致个体做出差异化选择的现象，即某个决策问题有两种描述方式不同但实质含义相同的语言表达形式，导致人们的选择结果出现显著差异（Tversky & Kahneman, 1981）。由于凸显的信息不同，信息框架效应有着不同的分类，既有反映事物特征信息的效价框架（例如某件事的利弊; Wong, 2020），也有反映事物量化信息的数字框架（例如某件事的发生概率; Chernev & Gal, 2010）。

过往的研究主要围绕获得-损失的效价框架展开（Levin et al., 1998）。根据前景理论，获得-损失框架通过强调决策目标的正向或负向特征，引起记忆中的有利或不利联想，最终造成个体对目标不同的认识（Kareklas et al., 2012）。典型的获得-损失框架效应如“牛肉问题”，即当分别以“含 25%的肥肉”和以“含 75%的瘦肉”描述同一个汉堡包，人们对前者产生更不健康的感知、更少的选择偏好和更弱的购买意愿（Levin et al., 1998; 文桂婵等, 2011）。Kim 和 Song（2022）曾调查在获得-损失框架下人们对 AI 的反应，对每 10 次任务成功 8 次且失败 2 次的 AI，研究者分别描述为“成功率 80%的 AI”和“失误率 20%的 AI”，受信息框架影响，人们给予“成功率 80%的 AI”更多的信任。

效价框架的应用虽然广泛，但用于表述“AI 失业”却不常见。一方面，“失业”一词容易引发人们的不利联想；另一方面，AI 失业往往涉及“数字”与“类别”信息，不易赋予效价。例如，“51%的岗位将被取代”（Chui et al., 2016），“证券、法律、保险、财会等行业属于被 AI 替代的高风险职业（Felten et al., 2021）”等。相比于强调获得与损失的效价特征，类似的表述方式更接近因素-概率的数字框架（Karmarkar & Kupor, 2023）。

因素框架是侧重强调目标事件的若干表现形式（Tversky & Koehler, 1994），例如相比于“AI 将在未来取代 77%的人类工作”的一般表述，使用因素框架后的表述为：“AI 将在未来取代包括服务业、制造业和农业在内 77%的人类工作”。概率框架则是在凸显各表现形式的概率细节（Baron, 2014），即“AI 将在未来取代包括服务业（24%）、制造业（26%）和农业（27%）在内 77%的人类工作”。在说服领域的研究中，因素框架具有提升风险评估等级的效果，当人们了解交通事故的若干类型后，系安全带行为的频率也会增加（Brewer et al., 2004）。在 Karmarkar 和 Kupor（2023）的研究中则发现，使用因素框架表述某种传染病能增加人们的预防行为，而使用概率框架表述同一疾病则导致相反的结果。

因素-概率框架之所以引起行为差异，是因为两种框架改变了人们对事件发生可能的主观判断（Morewedge & Kahneman, 2010）。在因素框架下，事件的各类表现形式均被人们

当作证明目标存在的额外佐证，累积的证据导致人们判断其发生的可能性大（Redelmeier et al., 1995; Rottenstreich & Tversky, 1997）。概率框架由于强调各表现形式的概率细节，导致人们以平均化的运算方式处理信息，各子概率平均后的结果便是发生的可能性低（Anderson, 1965; Chernev & Gal, 2010; Lynch, 1979; Weaver & Garcia, 2018）。换言之，因素框架激活了“求和”的计算过程，而概率框架激活了“求平均数”的计算过程，二者导致个体对事件发生的可能性做出不同的判断（Karmarkar & Kupor, 2023）。

因此在了解 AI 失业的有关内容时，分别接触因素框架或概率框架，能否导致人们在判断 AI 失业可能性方面的差异，而这种差异是否又会引发不同水平的 AI 威胁感？这是本研究想要回答的主要问题。

1.2 AI 威胁及对行为的影响

AI 威胁是人类心理威胁的一种具体形态（Gray et al., 2024; Złotowski et al., 2017）。根据社会认同理论（social identity theory），人们将自身归属于人类群体，与之相对是 AI 所归属的非人类群体，群体成员有维护所属群体的利益和独特性的动机，当群体的利益或独特性被外群体侵犯时，则引发群体冲突并造成威胁感（Brewer, 2001; Jackson, 1993; Tajfel & Turner, 1986）。按照 Riek 等人（2006）的分类，威胁分为现实威胁（realistic threat）和象征威胁（symbolic threat）：现实威胁危及个人所在群体的存续自保、经济利益或政治权利（Campbell, 1966），象征威胁则损害个人所在群体的身份特征或道德价值观（Stephan et al., 2015）。

Złotowski 等人（2017）对 AI 威胁的类型也进行了同样的划分，AI 现实威胁具体指 AI 可能给人类造成不安全后果以及对人类现有资源的争夺，如压缩人类生存空间与侵占人类经济利益等；AI 象征威胁指 AI 对人类智能的模仿导致人类与机器的身份界限趋于模糊，如 AI 展现出比拟人类的创造性，介入道德决策和宗教事务等（Samuel, 2020; Vincent, 2022）。多项研究结果均显示，人对 AI 的威胁感具有跨越地域、文化以及意识形态的普遍性（Dang & Liu, 2021; Gnambs & Appel, 2019; Yam et al.,2023）。

AI 威胁感上升通常导致人机关系恶化的结果，虽有调查发现，AI 的威胁感使员工产生更强的培训技能的意愿（Innocenti & Golin, 2022），但其目的在于避免工作被 AI 取代（Di Tella & Rodrik, 2020）。此外，更多的证据指向了负面的结果，例如个体当中出现的技术恐惧症（McClure, 2018）、AI 使用意向下降（Huang et al., 2021）、不能信任 AI（Lin et al., 2022）、以及反对将税款用于机器人研发的态度等等（Yogeeswaran et al., 2016）。具体到工作场景中，AI 威胁不仅加剧了员工所感受到的失业风险，也改变了员工对工作岗位的评价标准，并可能使其在职业选择中做出次优决策。例如，求职者更愿意选择工作内容侧重于人际交互和创造力的岗位（Rotman, 2013），技能需求较低且任务同质化严重的工作更难以进入个人的职业规划（Jørgensen & Navrbjerg 2001）。

需说明的是，医疗和经济领域所取得的许多成果证明了，人机协作的效果要优于人力

合作和全自动化办公，这种良性合作需要人们给予 AI 充分的信任（Tang et al., 2022; Waardenburg et al., 2022），然而 AI 威胁则会破坏信任的基础（Alaiad & Zhou, 2013; Correia et al., 2016; Paetzel et al., 2020）。如果因素与概率框架造成不同程度的 AI 威胁，那么较低的 AI 威胁感是否能够使人们对 AI 政策的研发持更加支持的态度（即 AI 曝光度高的工作）以及能否使个体更愿意选择（或推荐）那些需要经常接触 AI 的工作，这也是本研究需要探讨的问题。

1.3 模糊容忍度对信息框架效应的影响

信息框架效应体现了信息对人的引导，该效应的强弱也与个体特征相关联（Lauriola et al., 2005）。大五人格特质的外倾性、尽责性和神经质均影响个体受信息框架影响的程度以及因此产生的决策偏好（Levin et al., 2002; Rusting & Larsen, 1998）。个体的计算能力或对数学与概率理解力的差异，也反映为信息框架效应的差别，计算能力强的个体相比计算能力差的个体更容易识别数字中获取效价信息，因而受信息框架的影响更小（Peters et al., 2006; Stanovich & West, 1998）。另外，个人对目标事件的固有经验也会对决策中的认知加工产生一定的影响，对所描述的事物更为熟悉的个体不易受到信息框架影响，而缺少相关体验的个体往往由于决策情景的高度不确定性而产生信息框架效应（Hoch & Ha, 1986)）。最后，个体认知风格中的认知闭合需要也影响其受框架影响的程度，认知闭合需要水平较高的个体倾向于快速决策并对信息进行启发式加工，因而更易收到信息框架的影响（Kruglanski, 1989）。

AI 失业的特定背景下，一方面关于失业规模、强度和发展趋势存在诸多不确定性，个体在了解 AI 失业时所接收到的信息呈现出“模糊性”（Złotowski et al., 2017），另一方面，在工作实践的过程中，个体也不能穷尽所有职业信息之后再做决策，因此对于 AI 的威胁感知判断具有“及时性”（Xu & Tracey, 2014）。我们据此认为，基于 AI 失业的信息框架效应可能也与个人特质中的模糊容忍度有关。

模糊容忍度是个体对模糊情况的接受程度，反映一个人对新异、不一致、复杂的模糊不确定信息，进行容忍的特质倾向（Furnham & Ribchester, 1995），并与个体解决不确定性侵扰的迫切程度存在关联（Webster & Kruglanski, 1994）。模糊容忍度较低的个体在面临信息不完整的情境时，会感受到压力、并迅速地做出反应，以避免受到模糊性的刺激；而对模糊性容忍程度高的个体会将该情境感知为有趣，更接受复杂且不一致的的信息（Furnham & Ribchester, 1995）。当以不同信息框架表述“AI 失业”时，按照所含信息越少信息模糊度越高进行排序，则一般性表述最模糊（AI 将在未来取代 77%的人类工作），因素框架其次（AI 将在未来取代 77%的人类工作），概率框架又次之（AI 将在未来取代包括服务业（24%）、制造业（26%）和农业（27%）在内 77%的人类工作）。由于模糊容忍度低的个体表现出易受启发式加工影响、在不确定环境中倾向以极端化方式处理信息、对未来的过度焦虑等特点（Dugas et al., 1998; Grenier et al., 2005），对三种信息框架可能表

现出相似的反应。而模糊容忍度高的个体由于存在更多的自我探索及环境探索行为、有更高的生涯决策效能（Xu & Tracey, 2015），对不确定信息的细致加工可能造成其更易受到信息框架的影响。我们推测，在了解 AI 失业的相关内容时，信息框架效应更多存在于模糊容忍度较高的个体当中。

1.4 研究概览

综上所述，本研究通过 7 个实验考察因素-概率框架对 AI 威胁感知的影响。实验 1 发现，在了解 AI 失业的过程中，受概率框架（相比因素框架）影响的人们感知到的 AI 威胁更低。实验 2 考察了可能性的中介机制，概率框架影响下的人们认为 AI 失业发生的可能性更小导致他们感知 AI 的威胁更低。实验 3 排除了概率框架中具体数值大小的影响，复制了实验 1、2 的结果。实验 4 进一步检验了概率的信息框架效应，在概率框架的作用下，即便人们被告知的 AI 失业发生的预测规模更大，也依然感知 AI 威胁低。实验 5 检验了模糊容忍度的调节作用，概率框架导致人们感知 AI 的威胁更低，该效应更多出现在模糊容忍度高的个体中。实验 6、7 进一步探讨了这种框架效应对人们后续行为的影响，受概率框架影响，低水平的 AI 威胁感导致人们对 AI 技术的研发政策持更支持的态度，并对需经常接触 AI 的行业与职业有更强的推荐意愿。

2 实验 1：信息框架对 AI 威胁的直接影响

实验 1 的目的是考察信息框架对 AI 威胁感知的影响。本实验采用单因素三水平被试间的实验设计，被试间变量为信息框架类型，因变量分别为 AI 威胁感，以及对大规模 AI 失业发生可能性的判断。考察相比一般描述和因素框架，受概率框架影响的人是否认为大规模 AI 失业发生的可能性更低，以及对 AI 产生更低的威胁感。

2.1 被试

使用 G*Power 3.1（Faul et al., 2007）进行先验分析，设定效应量 $f = 0.2$，显著性水平 $\alpha = 0.05$，统计功效 power = 0.9，所需总样本量为 322。通过 Credamo 见数平台在线招募被试，最终样本量 $N = 322$（男性：136, 女性：186），年龄 18~67 岁（$M_{\text{age}} = 32.16$; $SD_{\text{age}} = 10.28$）。将所有被试随机分配到三种实验条件下，分别是一般条件组（$N = 111$）、因素条件组（$N = 106$）和概率条件组（$N = 105$），以小额现金作为报酬。

2.2 程序

所有被试首先要通过注意检查以确保了解工作分类的有关内容，以及我国的工作分类与其他多数国家是相似的。注意力检查的问题分别是："把人类的工作划分为物流运输、制造、安装维修、建筑冶炼、农业、办公室行政、销售、服务业、医疗卫生、科教文艺、计算机、金融商业管理等不同类型，是世界范围内普遍的准则，我们国家也不例外。为了确认您已经阅读了上述信息，请回答：我国的行业类型是否与大多数国家不同？"以及"对于大多数国家来说，农业、制造业和服务业是最常见的工作类型，我们国家也不例外。

为了确认您已经阅读了上述信息，请回答：在我国，农业、制造业和服务业是否和大多数国家一样都是常见的工作类型？”

通过注意检查后，所有被试会看到一条虚构的关于未来 AI 失业的测算结果：“预计 2033 年，AI 将替代 58%的人力岗位”。其中，“58%”的数值设定一方面参照了 Karmarkar 和 Kupor（2023）的研究方法，另一方面该数值介于中美两国“AI 失业”规模的真实测算结果之间，更能被被试所采信（美国为 51%，中国为 77%; Chui et al., 2016; 范长煜, 邓韵雪，2022）。

在一般条件组中，被试看到的补充说明为：“具体来说：有 58%的人的工作会被 AI 取代。”

在因素条件组中，被试看到的补充说明为：“有 58%的人的工作会被 AI 取代。具体来说：被 AI 取代的工作来自物流运输业；被 AI 取代的工作来自制造业；被 AI 取代的工作来自科教文艺行业；被 AI 取代的工作来自办公室行政；被 AI 取代的工作来自销售业；被 AI 取代的工作来自金融商业管理行业；此外，被 AI 取代的工作也来自其他行业。”

在概率条件组中，被试看到的补充说明为：“有 58%的人的工作会被 AI 取代。具体来说：被 AI 取代的工作约有 8%来自物流运输业；被 AI 取代的工作约有 7%来自制造业；被 AI 取代的工作约有 9%来自科教文艺行业；被 AI 取代的工作约有 8%来自办公室行政；被 AI 取代的工作约有 9%来自销售业；被 AI 取代的工作约有 7%来自金融商业管理行业；此外，被 AI 取代的工作约有 10%来自其他行业。”

在阅读材料之后，被试在“0 = 完全不可能”到“100 = 完全可能”的范围之内判断 AI 失业发生的可能性。被试接下来完成对 AI 威胁感知的测量，该量表采用李克特 7 点计分，包含 10 个条目，Cronbach's $\alpha = 0.93$，具体条目例如：“从长远来看，AI 技术对人类的安全和福祉构成直接威胁；AI 技术在日常生活中的日益普及对人类安全构成了威胁；技术的最新进步对人类的本质提出了挑战等（Złotowski et al., 2017）”。最后，所有被试报告了性别和年龄。

2.3 结果

以信息框架作为自变量，可能性感知和 AI 威胁感知作为因变量进行方差分析。结果显示，信息框架对 AI 威胁感知的总体影响显著，$F(2, 319)= 5.43$, $p = 0.005$, $\eta_p^2 = 0.03$, 95% CI [0.00, 0.08]，事后 Bonferroni 检验结果表明，概率条件组被试对 AI 威胁的感知低于单因素条件组（$p = 0.023$）和多因素条件组（$p =0.001$）。单因素条件组和多因素条件组之间在 AI 威胁感知上无显著差异（$p = 0.333$）。

信息框架对可能性的总体影响显著，$F(2, 319)= 11.05$, $p < 0.001$, $\eta_p^2 = 0.07$, 95% CI [0.02, 0.12]，事后检验 Bonferroni 结果表明，多概率条件组被试对 AI 取代人类工作的可能性感知低于单因素条件组（$p < 0.001$）和多因素条件组（$p < 0.001$）。单因素条件组和多因素条件组之间在 AI 取代人类工作的可能性感知上无显著差异（$p = 0.396$）（见表 1）。本实验

和后续实验在控制性别和年龄等变量后均不改变结果的显著性和方向。

表 1 不同信息框架条件威胁和可能性的平均值、标准差

因变量	信息框架	*M*	*SD*
威胁感知	多概率条件组	4.29	1.38
	单因素条件组	4.69	1.23
	多因素条件组	4.85	1.18
可能性感知	多概率条件组	53.26	17.69
	单因素条件组	60.49	13.15
	多因素条件组	62.18	12.71

2.4 讨论

实验 1 检验了信息框架在 AI 威胁以及 AI 失业发生可能性上所造成的差异。结果显示，当人们看到对 AI 失业的规模预测（58%）后，相比于一般性描述和因素框架描述，受到概率框架影响，即使各行业的具体概率之和（58%）与预测结果一致，人们会认为 AI 的威胁更低、AI 失业发生的可能性更小，从而验证了假设 1。

由于一般条件组和因素条件组在 AI 威胁与可能性感知上均不存在显著差异，信息框架的影响作用主要体现在概率条件下。而相比于一般条件组，因素条件与与概率条件所含的信息量更大，因此在后续的实验中，我们将主要的比较对象设定为因素条件组和概率条件组。以上结果扩展了概率-因素的数字信息框架效应的结果（Karmarkar & Kupor, 2023），表明在个体了解 AI 失业的过程中，信息框架对可能性感知和 AI 威胁感知均存在影响。

3 实验 2：可能性感知的中介作用

由于实验 1 的结果表明，受概率框架影响，人们判断 AI 失业发生的可能性更低，对 AI 的威胁感知更弱。因此，实验 2 的目的便在于考察信息框架对 AI 威胁的差异是否源自于感知可能性的差异，即当人们主观上认为工作被 AI 取代的可能性越小，对 AI 的威胁感知便越低。

3.1 被试

实验 2 采用单因素两水平被试间的实验设计，使用 G*Power 3.1（Faul et al., 2007）进行先验分析, 设定效应量 $f = 0.2$, 显著性水平 $\alpha = 0.05$,统计功效 power = 0.9, 所需总样本量为 266。通过 Credamo 见数平台在线招募被试，最终样本量 $N = 266$（男性：114, 女性：152），年龄 18~67 岁（$M_{age} = 33.18$; $SD_{age} = 11.57$），将所有被试随机分配到两种实验条件下，分别是因素条件组（$N = 135$）和概率条件组（$N = 131$）。

3.2 程序

与实验 1 一致，被试首先通过注意检查。在看过对未来 AI 引发失业规模的预测之后，

被试阅读了因素条件或概率条件下的补充说明。可能性感知与 AI 威胁的测量与实验 1 相同（Cronbach's α = 0.94）。考虑到不同被试对算法的看法和知识可能存在差异，从而影响到其对 AI 威胁感知，为了排除相关可能的影响因素，被试还被要求报告他们对算法的熟悉程度（“你对算法有多熟悉？”）、了解程度（“与普通中国人相比，你认为你对算法有多了解？”）和喜爱程度（“你有多喜欢算法?）（Bartneck et al., 2009; Leo & Huh, 2020; 许丽颖 等, 2021）。最后，所有被试报告了性别和年龄。

3.3 结果

对两种信息框架下被试的 AI 威胁进行方差分析，结果显示，概率条件组被试（M = 4.27; SD = 1.38）的 AI 威胁感知低于因素条件组（M = 4.86; SD = 1.33），$p < 0.001$。该结果再次验证了假设 1。

以信息框架类型为自变量（因素条件编码为 0，概率条件编码为 1），AI 失业的可能性感知为中介变量，AI 威胁为因变量，在样本选择为 5000 次、95%的置信区间下采用 Bootstrap（PROCESS Model 4; Hayes, 2013）分析中介效应。结果显示，可能性感知的中介效应显著（indirect effect b = –0.27, SE = 0.07, 95% CI = [–0.40, –0.15]）。如图 1 所示，比起因素条件组，受概率条件影响的人们认为 AI 失业发生的可能性更小，因此对 AI 的威胁更低，该结果验证了可能性感知在信息框架影响 AI 威胁的关系中发挥中介作用的假设。

此外，在控制了被试对 AI 的熟悉度之后，AI 失业的可能性感知在信息框架与 AI 威胁之间的中介效应仍然显著（indirect effect b = –0.25, SE = 0.06, 95% CI = [–0.38, –0.14]）。

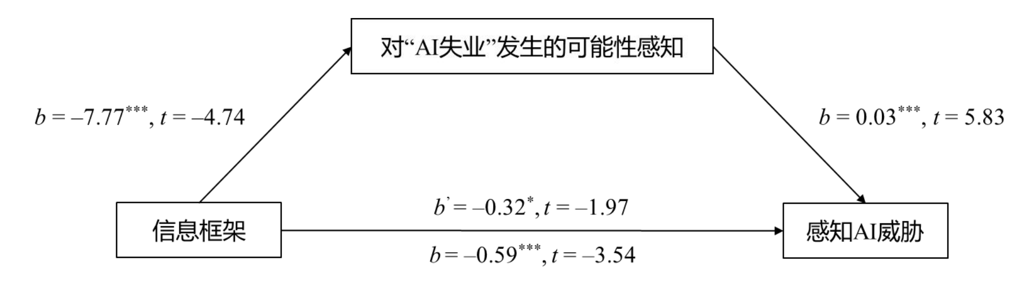


图 1 对“AI 失业”发生的可能性感知的中介作用

（信息框架：0 = 因素框架，1 = 概率框架；* $p < .05$. ** $p < .01$. *** $p < .001$）

3.4 讨论

实验 2 通过中介检验，验证了感知可能性的中介作用。结果表明，比起因素框架，概率框架影响下的人们认为 AI 失业发生的可能性更小，进而对 AI 产生更低的威胁感。这一发现进一步表明信息框架对主观可能性感知的重要影响（Eagly & Chaiken, 1993）。在后续的实验中，我们将变更实验条件继续考察可能性中介作用的稳定性。

4 实验 3：排除概率数值的潜在影响

实验 3 的目的有两个。首先，实验 3 要排除概率数值对主观可能性的影响，由于在此前实验中，概率条件组 7 个行业里被 AI 取代的岗位之和虽然是 58%，但具体到每个行业的

AI 失业规模则为 7%~10%之间，其中存在一种可能性，即因为每个行业对应的概率值偏低导致人们认为 AI 失业更不可能发生，因此，我们希望排除通过调整实验材料来排除这种替代性假设。其次，实验 3 再次加入对 AI 失业发生主观可能性的测量，来继续检验可能性感知对于信息框架影响 AI 威胁的中介作用。

4.1 被试

实验 3 采用单因素两水平被试间的实验设计，使用 G*Power 3.1（Faul et al., 2007）进行先验分析, 设定效应量 $f = 0.2$, 显著性水平$\alpha = 0.05$,统计功效 power = 0.9, 所需总样本量为 266。通过 Credamo 见数平台在线招募被试，最终样本量 $N = 266$（男性：108, 女性：158），年龄 19~57 岁（$M_{age} = 29.66$; $SD_{age} = 7.59$），将所有被试随机分配到两种实验条件下，分别是因素条件组（$N = 133$）和概率条件组（$N = 133$）。

4.2 程序

在通过注意检查，确保已了解人力工作的基本类型之后，所有被试看到针对未来国内 AI 失业规模的预测结果：“预计 2033 年，AI 将替代 58%的人力岗位”。

接下来，因素条件组的被试看到的补充说明为：“具体来说，被取代工作岗位来自服务业；被取代的工作岗位也来自服务业以外的其他行业”。

概率条件组被试看到的补充说明为：“具体来说，被取代的工作岗位当中有 28%来自服务业；被取代的工作岗位当中有 30%来自服务业以外的其他行业”。

之后，被试与此前实验一样，完成对 AI 失业的可能性判断以及对 AI 威胁的测量（Cronbach's $\alpha = 0.92$）。最后，所有被试报告了性别和年龄。

4.3 结果

对两种信息框架下被试感知到的 AI 威胁进行方差分析，结果显示，概率条件组被试（$M = 4.24$; $SD = 1.23$）的 AI 威胁感知低于因素条件组（$M = 4.71$; $SD = 1.16$），$p < 0.001$。该结果同样验证了信息框架影响 AI 威胁的假设。

以信息框架类型为自变量（因素条件编码为 0，概率条件编码为 1），AI 失业的可能性感知为中介变量，AI 威胁为因变量，在样本选择为 5000 次、95%的置信区间下采用 Bootstrap（PROCESS Model 4; Hayes, 2013）分析中介效应。结果显示，可能性感知的中介效应显著（indirect effect $b = –0.15$, $SE = 0.06$, 95% CI = [–0.27, –0.05]。比起因素条件组，受概率条件影响的人们认为 AI 失业发生的可能性更小，因此对 AI 的威胁更低，该结果再次验证了可能性感知在信息框架影响 AI 威胁的关系中发挥中介作用的假设。

4.4 讨论

综合以上结果，实验 3 再次验证了信息框架对于 AI 威胁感知的影响，尤其是当对 AI 失业的补充说明由 7 类行业降至 2 类行业、各行业对应的失业概率由 7%~10%上升至 28%和 30%，受概率框架影响的人们依然会对 AI 产生更低的威胁。此外，实验 3 的结果显示，

感知可能性的中介作用仍然存在，框架条件组的被试在主观上认为AI失业发生的可能性更小，进而认为AI的威胁更低。

5 实验4：排除失业规模的潜在影响

实验4希望再次验证信息框架对AI威胁影响作用的稳健性，具体而言，我们希望能够在重复实验3结果的基础上提出进一步的假设：相比于因素框架条件组，受概率框架影响的被试即便阅读到AI失业规模更大的预测结果，也仍然会形成AI失业更不可能发生的主观判断。

5.1 被试

实验4采用单因素两水平被试间的实验设计，使用G*Power 3.1（Faul et al., 2007）进行先验分析，设定效应量 $f = 0.2$，显著性水平 $\alpha = 0.05$,统计功效 power = 0.9, 所需总样本量为266。通过Credamo见数平台在线招募被试，最终样本量 $N = 268$（男性：117, 女性：151），年龄18~59岁（$M_{age} = 30.29$; $SD_{age} = 8.21$），将所有被试随机分配到两种实验条件下，分别是因素条件组（$N = 134$）和概率条件组（$N = 134$）。

5.2 程序

与此前实验不同，在通过注意检查之后，因素条件组的被试看到针对未来国内AI失业的预测结果为："预计2033年，AI将替代**58%**的人力岗位"。而概率条件组被试则看到："预计2033年，AI将替代**59%**的人力岗位"。两组被试看到的预测规模并不相同。

接下来，因素条件组的被试看到的补充说明为："具体来说，被取代工作岗位来自服务业；被取代的工作岗位也来自服务业以外的其他行业"。概率条件组被试看到的补充说明为："具体来说，被取代的工作岗位当中有29%来自服务业；被取代的工作岗位当中有30%来自服务业以外的其他行业"。之后，被试完成对AI失业的可能性判断以及对AI威胁的测量（Cronbach's $\alpha = 0.92$）。最后，所有被试报告了性别和年龄。

5.3 结果

对两组被试的AI威胁感知进行方差分析，结果显示，概率条件组被试（$M = 4.19$; $SD = 1.21$）的AI威胁感知低于因素条件组（$M = 4.65$; $SD = 1.24$），$p = 0.002$。对两组被试的可能性判断进行方差分析，结果显示，概率条件组被试（$M = 56.82$; $SD = 15.45$）认为AI失业发生的可能性低于因素条件组（$M = 61.96$; $SD = 12.35$），$p = 0.003$。上述结果再次验证了信息框架影响AI威胁以及可能性感知的假设。

以信息框架类型为自变量（因素条件编码为0，概率条件编码为1），AI失业的可能性感知为中介变量，AI威胁为因变量，在样本选择为5000次、95%的置信区间下采用Bootstrap（PROCESS Model 4; Hayes, 2013）分析中介效应。结果显示，可能性感知的中介效应显著（indirect effect $b = -0.08$, $SE = 0.04$, 95% CI = [−0.17, −0.02]），该结果也再次验证了可能性感知在信息框架与AI威胁关系中发挥中介作用的假设。

5.4 讨论

实验 4 再次检验了信息框架对 AI 威胁的影响，并重复了实验 3 的结果，即当概率框架的 AI 失业规模由 7%~10%上升至 28%~30%，该组被试仍然对 AI 产生较低的威胁感。更为重要的是，通过实验 4 我们发现，即使概率条件组看到的 AI 失业规模（59%）大于因素条件组（58%），还是会认为 AI 失业发生的可能性更小，以及认为 AI 的威胁更低。实验 4 也再次验证了感知可能性的中介作用，即框架条件组的被试在主观上认为 AI 失业发生的可能性更小，进而认为 AI 的威胁更低。总体而言，通过实验 1~4，我们检验了信息框架、AI 威胁、可能性感知三者关系的稳健性，后续实验将开始探讨信息框架对 AI 威胁影响成立的边界条件。

6 实验 5：模糊容忍度的调节作用

实验 5 的目的在于考察模糊容忍度是否在信息框架与 AI 威胁的关系中发挥调节作用。理论上模糊容忍度低的个体在接收信息后，易受到启发式加工的影响、难以进行全面的分析、也难以做出审慎判断。因此，信息框架对于 AI 威胁的影响是否只发生在模糊容忍度较高的个体身上，这是实验 5 想要回答的问题。

6.1 被试

实验 5 采用单因素两水平被试间的实验设计，使用 G*Power 3.1（Faul et al., 2007）进行先验分析，设定效应量 $f = 0.2$，显著性水平 $\alpha = 0.05$，统计功效 power = 0.9，所需总样本量为 266。通过 Credamo 见数平台在线招募被试，最终样本量 $N = 264$（男性：93, 女性：171），年龄 19~51 岁（$M_{age} = 28.80$; $SD_{age} = 7.00$），将所有被试随机分配到两种实验条件下，分别是因素条件组（$N = 138$）和概率条件组（$N = 126$）。

6.2 程序

与实验 2 相同，通过注意检查后，所有被试看到针对未来国内 AI 失业的预测结果。接下来两组被试分别看到不同的补充说明：因素条件组的补充材料表明 AI 失业发生在 7 类行业当中；概率条件组的补充材料既包含了 AI 失业发生的 7 类行业，也包括了每个行业对应的概率信息。被试随后完成对 AI 失业的可能性判断以及对 AI 威胁的测量（Cronbach's α = 0.91）。然后，我们使用中文版多类型刺激模糊容忍度量表（Multiple Stimulus Types Ambiguity Tolerance Scale–II，简称 MSTAT–II; McLain, 2009）进行测量，该量表共 13 个条目，分别对应新异性（如：“相比于新环境，我更喜欢熟悉的环境”）、复杂性（如：“我喜欢解决那些信息复杂，线索不清的问题”）、无解性（如：“我不愿解决需要从不同角度进行分析的问题”）和不确定性（如：“我发现在结果不确定时候，很难做出选择”）等 4 个维度（Cronbach's α = 0.87）。最后，所有被试报告了性别和年龄。

6.3 结果

对两组被试的 AI 威胁感知进行方差分析，结果显示，概率条件组被试（$M = 4.21$; SD

= 1.16）的 AI 威胁感知低于因素条件组（M = 4.60; SD = 0.99），p = 0.004。对两组被试的可能性判断进行方差分析，结果显示，概率条件组被试（M = 56.42; SD = 18.15）认为 AI 失业发生的可能性低于因素条件组（M = 61.25; SD = 15.66），p = 0.021。

以信息框架类型为自变量（因素条件编码为 0，概率条件编码为 1），AI 失业的可能性感知为中介变量，AI 威胁为因变量，在样本选择为 5000 次、95%的置信区间下采用 Bootstrap（PROCESS Model 4; Hayes, 2013）分析中介效应。结果显示，可能性感知的中介效应显著（indirect effect b = –0.12, SE = 0.05, 95% CI = [–0.23, –0.02]）。

以 AI 威胁为因变量考察信息框架（因素条件编码为 0，概率条件编码为 1）与模糊容忍度的交互作用，结果表明信息框架和模糊容忍度对 AI 威胁存在显著的交互作用（b = –0.44, SE = 0.20, t = –2.19, p = 0.029），概率条件组的 AI 威胁感显著低于因素条件组（b = –0.39, SE = 0.12, t = –3.15, p = 0.002），模糊容忍度对 AI 威胁存在显著影响（b = –0.65, SE = 0.10, t = –6.51, $p < 0.001$），模型的调整 R^2 = 018, ΔR^2 = 0.02, F = 4.80, p = 0.03。交互作用如图 2 所示，简单斜率分析结果表明，对于模糊容忍度较低的个体而言，信息框架对 AI 威胁的影响不显著（b = –0.12, SE = 0.17, t = –0.67, p = 0.50）；对于模糊容忍度较高的个体而言，信息框架对 AI 威胁的影响显著（b = –0.66, SE = 0.17, t = –3.78, $p < 0.001$）。

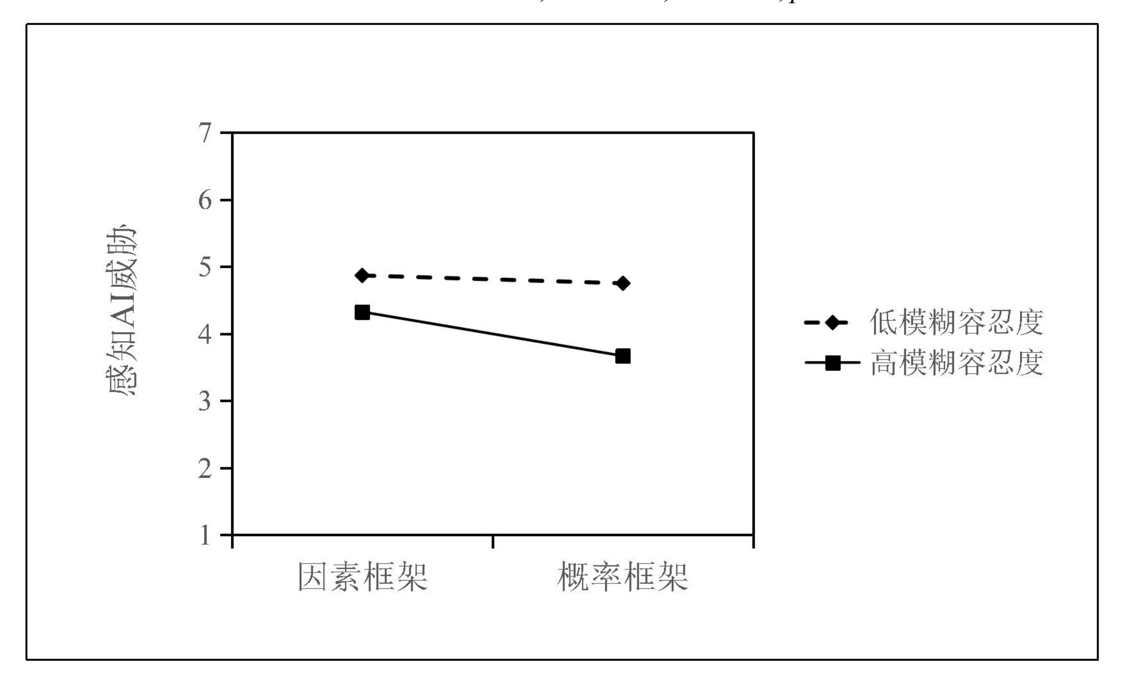


图 2 模糊容忍度的调节作用

6.4 讨论

实验 5 进一步探索了信息框架影响 AI 威胁的边界条件，结果发现模糊容忍度在信息框架对 AI 威胁的影响中发挥调节作用。在模糊容忍度较高的个体当中，受概率框架的影响的人们感受到的 AI 威胁比因素条件组更低；而在模糊容忍度较低的个体当中，两种信息框架组感受到的 AI 威胁无显著差异。同时，实验 5 还再次验证了可能性感知的中介作用，即由于概率框架降低对 AI 失业发生可能的判断，进而导致感受到更低的 AI 威胁。在后续的实

验中，我们将在信息框架影响 AI 威胁的基础上，继续探讨 AI 威胁感的差异又会导致人们在态度、行为层面产生怎样的变化。

7 实验 6：信息框架对 AI 研发的支持

实验 6 探究的是信息框架对 AI 威胁造成影响后所引发的行为层面的变化。在 Złotowski 及其同事（2017）的研究中发现，感知到更高的 AI 威胁会降低人们对于 AI 研发相关政策的支持，因此我们想要检验信息框架也会通过对 AI 威胁的影响造成人们在 AI 研发政策支持度上的差异。

7.1 被试

实验 6 采用 2（信息框架：概率框架 vs. 因素框架）× 2（行业类型：材料涉及的行业 vs. 材料未涉及的行业）的混合设计（其中，被试间因子为信息框架，被试内因子为行业类型），使用 G*Power 3.1（Faul et al., 2007）进行先验分析，设定效应量 $f = 0.2$，显著性水平 $\alpha = 0.05$，统计功效 power = 0.9，所需总样本量为 266。通过 Credamo 见数平台在线招募被试，最终样本量 N = 261（男性：110, 女性：151），年龄 18~58 岁（$M_{age} = 29.17$; $SD_{age} = 7.85$），将所有被试随机分配到两种实验条件下，分别是因素条件组（$N = 129$）和概率条件组（$N = 132$）。

7.2 程序

与实验 2、实验 5 相同，所有被试通过注意检查的被试在看过对 AI 失业的预测结果后，因素条件组和概率条件组的被试分别阅读了不同的补充说明，并对感知到的 AI 威胁进行测量（Cronbach's $\alpha = 0.87$）。此后，我们收集了被试对 AI 研发政策的支持程度，所使用的量表由 5 个条目构成，例如“你在多大程度上支持 AI 研究”、“你在多大程度上支持国家划拨经费投入 AI 研究”等（Złotowski et al., 2017）。

此外，为了解信息框架对 AI 威胁及政策支持度的影响是否仅限于实验材料涉及的行业中，我们还询问了被试两个问题：“你在多大程度上支持在物流、法律、投资、投资、行政、保险等前文涉及的行业中推进人工智能研究？”以及“你在多大程度上支持在建筑、手工、餐饮、文艺、纺织等前文没有涉及的行业中推进人工智能研究？”最后，所有被试报告了性别和年龄。

7.3 结果

对两组被试的 AI 威胁感知、AI 研发的政策支持度、在材料涉及的行业推进 AI 研发的支持度、以及在材料未涉及的行业推进 AI 研发政策的支持度进行方差分析。结果显示，概率条件组被试（$M = 3.95$; $SD = 0.98$）的 AI 威胁感知低于因素条件组（$M = 4.59$; $SD = 0.87$），$p < 0.001$；概率条件组被试（$M = 5.25$; $SD = 1.01$）对 AI 研发的政策支持度高于因素条件组（$M = 4.68$; $SD = 1.23$），$p < 0.001$；在材料中所涉及的 7 类行业中，概率条件组被试（$M = 5.03$; $SD = 1.15$）对在其中推进 AI 研发的支持度高于因素条件组（$M = 4.61$;

SD = 1.23），*p* = 0.005；而在材料中未涉及的 7 类行业中，概率条件组被试（*M* = 4.80; *SD* = 1.42）对在其中推进 AI 研发的支持度也高于因素条件组（*M* = 4.36; *SD* = 1.54），*p* = 0.018。

以信息框架类型为自变量（因素条件编码为 0，概率条件编码为 1），AI 威胁为中介变量，对 AI 研发政策的支持度为因变量，在样本选择为 5000 次、95%的置信区间下采用 Bootstrap（PROCESS Model 4; Hayes, 2013）分析中介效应。结果显示，AI 威胁的中介效应显著（indirect effect *b* = 0.13, *SE* = 0.06, 95% CI = [0.03, 0.26]），另见图 3。

以信息框架类型为自变量（因素条件编码为 0，概率条件编码为 1），AI 威胁为中介变量，以在材料涉及的行业推进 AI 研发的支持度为因变量，在样本选择为 5000 次、95%的置信区间下采用 Bootstrap（PROCESS Model 4; Hayes, 2013）分析中介效应。结果显示，AI 威胁的中介效应显著（indirect effect *b* = 0.14, *SE* = 0.06, 95% CI = [0.03, 0.27]）。

以信息框架类型为自变量（因素条件编码为 0，概率条件编码为 1），AI 威胁为中介变量，以在材料未涉及的行业推进 AI 研发的支持度为因变量，在样本选择为 5000 次、95%的置信区间下采用 Bootstrap（PROCESS Model 4; Hayes, 2013）分析中介效应。结果显示，AI 威胁的中介效应不显著（indirect effect *b* = 0.06, *SE* = 0.07, 95% CI = [–0.08, 0.21]）。

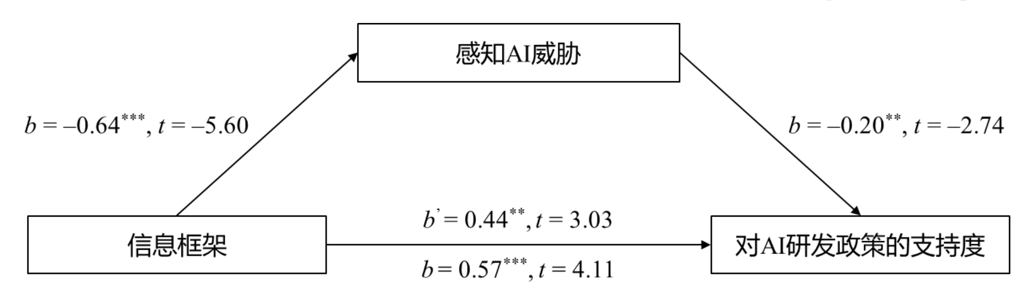


图 3. 感知 AI 威胁的中介作用

（信息框架：0 = 因素框架，1 = 概率框架；* *p* < .05. ** *p* < .01. *** *p* < .001）

7.4 讨论

实验 6 在信息框架与 AI 威胁之间关系的基础上继续考察个体由此在态度层面出现的差异，结果发现，相比于受到因素框架影响的人们，受概率框架影响的人们对推进 AI 研发的相关政策的支持度更高，更为重要的是，这种对 AI 研发政策的支持态度仅体现在补充材料涉及的行业中。此外，我们也发现信息框架对政策支持度的影响是通过对 AI 威胁的感知实现的，具体而言，是由于概率框架降低了人们对 AI 的威胁感知，进而导致人们对（在不同行业中）开展 AI 研发的政策更加支持。实验 6 将信息框架所造成的影响由个体的威胁感知扩展到对相关政策的态度上，在最后一个实验中，我们将继续探讨信息框架造成的后续影响是否也发生现实情景当中。

8 实验 7：信息框架对工作推荐意愿的影响

实验 7 共有四个目的。第一，想要考察在给予他人就业建议的场景当中，以不同的信息框架描述 AI 失业是否会引发人们对就业推荐度的变化。第二，由于在实验 6 中发现，信息框架对研发 AI 的政策支持度的影响会扩散到材料未涉及的行业中，因此实验 7 想要检验信息框架对就业建议的影响是否也存在类似的扩散效应。第三，考虑到不同职业在被 AI 替代的风险上存在差异，实验 7 的目的也在于考察信息框架以及职业自身可替代风险对于就业建议的综合影响。最后，我们还要验证信息框架是通过 AI 威胁对就业建议造成影响的。

8.1 被试

实验 7 采用 2（信息框架：概率框架 vs. 因素框架）× 2（职业的替代风险：高替代风险职业 vs. 低替代风险职业）的混合设计（其中，被试间因子为信息框架，被试内因子为职业的替代风险），使用 G*Power 3.1（Faul et al., 2007）进行先验分析，设定效应量 $f = 0.2$，显著性水平 $\alpha = 0.05$，统计功效 power = 0.9，所需总样本量为 266。通过 Credamo 见数平台在线招募被试，最终样本量 $N = 260$（男性：83, 女性：177），年龄 18~66 岁（$M_{age} = 29.64$; $SD_{age} = 8.49$），将所有被试随机分配到两种实验条件下，分别是因素条件组（$N = 130$）和概率条件组（$N = 130$）。

8.2 程序

在通过注意力检查后，所有被试阅读对 AI 失业的预测结果，因素条件组和概率条件组的被试分别阅读了不同的补充说明。接下来，在一个虚构的亲友求职场景下，我们询问了被试针对不对职业的推荐程度："假设一位亲戚朋友正面临找工作的问题，并向你征询就业方向的意见。暂不考虑其专业背景，并且各岗位间的工作量、待遇等基本相当。结合材料，你对他（她）将要从事材的职业会给出什么意见？（1 = 非常不推荐；7 = 非常推荐）"

用以评价的职业来自 Felten 及其同事（2021）对职业的 AI 曝光度的调查研究，被试看到的职业共有 10 种，分别属于高曝光（具体包括：行政文员、采购代购、信用审核、保险销售、中介咨询）和低曝光（具体包括：工程建筑、装修设计、文化宣传、环境规划、健身健美）两个类别。在完成职业推荐之后，被试完成 AI 威胁的测量（Cronbach's $\alpha = 0.88$），并报告了性别和年龄。

8.3 结果

对两组被试的 AI 威胁感知、对低 AI 曝光的职业推荐度、对高 AI 曝光职业的推荐度进行方差分析。结果显示，概率条件组被试（$M = 3.93$; $SD = 1.05$）的 AI 威胁感知低于因素条件组（$M = 4.39$; $SD = 1.03$），$p < 0.001$；对于低 AI 曝光的职业，概率条件组被试（$M = 4.85$; $SD = 0.87$）的推荐程度低于因素条件组（$M = 5.08$; $SD = 0.77$），$p = 0.022$；对于高 AI 曝光的职业，概率条件组被试（$M = 3.70$; $SD = 0.99$）的推荐程度高于因素条件组（$M = 3.17$; $SD = 0.98$），$p < 0.001$。

进一步的重复测量方差分析结果显示，信息框架与职业的 AI 曝光度之间存在交互作用（$F = 24.03, p < 0.001, \eta_p^2 = 0.09$）；如图所示，对于 AI 曝光度低的职业，概率条件组被试

（M = 4.85; SD = 0.87）的推荐意愿低于因素条件组（M = 5.08; SD = 0.77 , F = 19.20, p < 0.001, $\eta_p^2 = 0.07$）；对于 AI 曝光度高的职业，概率条件组被试（M = 3.70; SD = 0.99）的推荐意愿高于因素条件组（M = 3.17; SD = 0.98, F = 5.30, p = 0.022, $\eta_p^2 = 0.02$），另见图 4。

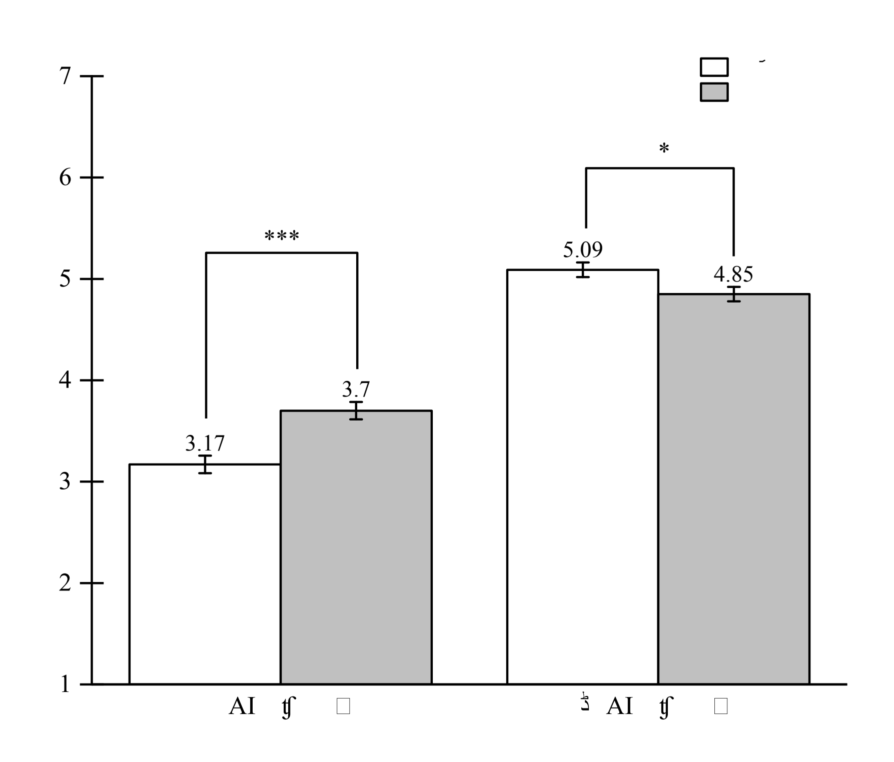


图 4 信息框架与 AI 曝光度对职业推荐意愿的交互作用

（* p < .05. ** p < .01. *** p < .001）

8.4 讨论

实验 7 在向他人提供建议的场景下探索了信息框架对于 AI 威胁、AI 曝光度不同的职业推荐意愿的影响。结果显示，相比于受到因素框架影响的人们，受概率框架影响的人们感受到更低的 AI 威胁。然而，在面对不同的职业时，信息框架对推荐程度的影响却表现出相反的趋势：相比因素框架，受概率框架影响的人们对于 AI 曝光度高职业的推荐程度更高，而对 AI 曝光度低的职业的推荐程度更低。另外，AI 威胁没有在其中发挥显著的中介作用，即信息框架对不同类型职业推荐度的影响并不是通过人们所感知到的 AI 威胁来实现的。

9 总讨论

本研究区分了关于 AI 失业的两种信息框架，系统考察了关于 AI 威胁以及就业决策的信息框架效应，并进一步考察了对大规模 AI 失业发生可能性的判断以及模糊容忍度在其中的作用。研究发现，受概率框架影响的个体所产生的 AI 威胁感更低（实验 1–7），对 AI 技术的研发政策更加支持（实验 6），并对 AI 曝光度高的职业有更强的推荐意愿（实验

7）。对 AI 失业发生的可能性判断在信息框架与 AI 威胁感的关系中起中介作用，受概率框架影响后，个体得出 AI 失业发生可能更低的主观判断，进而产生较低的 AI 威胁（实验 2–5）。模糊容忍度在信息框架与 AI 威胁感的关系中发挥调节作用，对于模糊容忍度较高者，概率框架会降低 AI 威胁感；但对于模糊容忍度较低者，这种影响则不显著（实验 5）。

9.1 信息框架对 AI 威胁的影响

本研究的主要结果证明了人们在了解“AI 失业”的过程中存在的信息框架效应，表现为概率与因素两种数字信息框架会造成不同程度的 AI 威胁感。依据前景理论，框架的本质在于选取并强调信息，强调导致人们从不同的角度看待问题（Kahneman & Tversky, 1979）。此前的研究通常采用获得-损失框架，通过对同一目标的积极特征与消极特征的信息处理，影响个体对于事物的态度、决策和行为（Mcelroy & Seta, 2003）。并得到以获得-损失框架表述 AI 技术，导致对 AI 的信任程度存在差异的结论（Kim & Song, 2022）。但由于民众对 AI 失业的认识往往来自 AI 是否、何时以及在何种规模上取代人类劳动的数字信息，且“失业”一词易引发负面联想，我们认为 AI 失业的相关信息可能涉及因素与概率两种信息框架（Acemoglu & Restrepo, 2020; Koch et al., 2021; Chui et al., 2016; Shoss & Ciarlante, 2022）。

在此基础上我们发现了相比因素框架，以概率框架表述 AI 失业导致人们产生更低的 AI 威胁、更加支持 AI 技术的研发政策、对于要经常接触 AI 的工作岗位也有更强烈的推荐意愿。Karmarkar 与 Kupor（2023）通过研究发现，呈现出流行病的若干传染源，人们对此的风险评估等级更高，而呈现每种传染源的概率细节，人们的风险评估等级便会下降。如果以风险评估的视角看待 AI 对人的威胁（生存资源和身份归属方面; Gray et al., 2024; Złotowski et al., 2017），则本文取得了与前人相一致的结论，能够完善信息框架效应领域的研究成果。

人机关系随着 AI 技术的进步趋于复杂，对 AI 失业的探讨不仅关乎 AI 技术的替代效应和创造效应最终造成了何种结果（Harari, 2017），也与人们所接触的信息、所做出的反应密切相关（Yam et al., 2023）。AI 失业的不确定性反映了人们从无到有、由寡及多地认识事物的过程（Blanas et al., 2019; Dauth et al., 2018），而其中关于 AI 失业的调查、预测、报道和观点均可能源自整体事件的某些特定方面（Gray et al., 2024），并引发不同类型的信息框架效应（Tversky & Kahneman, 1981）。本研究检验了数字信息框架所造成的影响，尤其关注 AI 威胁感，其启示意义在于，当 AI 失业成为舆论焦点时，对于宣传报道的内容等外部因素应慎重对待，因为信息的选择影响人们对此的态度，而态度影响着 AI 技术的应用。

9.2 概率-因素框架对主观可能性判断的影响

获得-损失框架效应之所以能引起决策与行为的差异，是因为获得框架引发了个体的积极联想而损失框架引发了消极联想，对效价信息加工的不同引发了框架效应（Kareklas et

al., 2012）。与之不同，因素-概率框架效应形成的机制在于主观可能性的变化：受到因素框架影响，个体以“相加求和”的方式去理解目标，往往得出发生可能性更大的结论；而受到概率框架影响，个体以“求平均数”的方式进行理解，则得出可能性更小的结果（Karmarkar & Kupor, 2023）。

人们从多个领域接收到多种形式的风险相关信息，在与失业有关的情况下，解释此类信息显得尤其重要。人们经常会遇到一系列的信息，他们对风险可能性的主观感知经常会引导他们决定是否采取防御措施（Brewer et al., 2004; Diener & Richardson, 2007; Meadows et al., 1993）。此前心理学和行为经济学等领域的研究表明，人在风险评估的过程中存在多种偏差，在本研究中，受特定信息引导而改变对 AI 失业发生可能的主观判断也是认知偏差的一种体现，即提供 AI 失业的额外概率细节会使人认为其更可能发小，从而改变人们做出相应决策的方式。

需要说明的是，实验 1 并没有得出因素框架相比一般性表述增加人们感知 AI 失业发生可能性的结论。对该结果的解释可能包括两个原因：首先是在两组参与者所接触的实验材料中（预计 2033 年，AI 将替代 58%的人力岗位; 龚遥, 彭希哲, 2020; 李磊, 何艳辉, 2019 ），58%作为唯一的数字信息，使参与者在判断“AI 失业”的可能性时产生了锚定效应；二是实验的参与者已经接触过关于 ChatGPT 等大语言模型的报道，在因素框架组所呈现的实验材料并非其判断 AI 失业发生可能性的唯一信息来源（Kahn, 2023）。在此之外，概率框架相比于因素框架导致主观可能性判断更低的效应是稳定的，在排除个体对 AI 熟悉程度影响的基础上，该效应也不受 AI 失业涉及的具体行业数量和预测规模数值大小的影响。

9.3 模糊容忍度的调节作用

模糊容忍度是个体在面对陌生复杂、存在不确定性或存在矛盾冲突的刺激时，从不接受到接受的范围变量（Mclain, 1993），它对应个体是否将不确定的状况视为一种极端状况，以及以何种方式应对不确定的侵扰（Grenier et al., 2005）。另外，模糊容忍度也决定了人在面对难辨真伪的信息时，是否能够接受并做出理性的决策（Katsaros & Nicolaidis, 2012）。目前有关 AI 失业的理论预测和调查实证均存在大量的探讨空间，这些争议和不一致的结论造成了诸多的不确定性（Acemoglu & Restrepo, 2020; Koch et al., 2021; Chui et al., 2016），因此，我们认为人们了解 AI 失业的过程，以及看待 AI 失业的态度，都与其对模糊信息的容忍度存在关联。

本研究发现，AI 失业的信息框架效应存在于模糊容忍度较高的个体当中，模糊容忍度水平高的个体往往表现出更多地探索行为，虽然此类个体收集的信息更多，对目标的评价更全面，做出的决策也更为理性，但也更易被信息框架所引导（Xu & Tracey, 2014）。与之相反，模糊容忍度低的个体的思维方式更加僵化刻板、在复杂的状况下往往过早选择并坚持解决方案的单一性，不承认事物具有好坏两面特征，寻求确定性，并倾向于逃避复杂情境（Frenkel-Brunswick, 1949），模糊容忍度低的个体也会对未来产生过度的担心，伴随

着更严重的焦虑情绪，最终导致更为消极的职业决策结果（Buhr & Dugas, 2006; Dugas et al., 1998; Dugas rt al., 2001; Keenan, 1978），在诸多因素的作用下，低模糊容忍度者倾向于以单一、极端的方式应对 AI 失业，信息框架效应反而难以体现（Frenkel-Brunswick, 1949）。

9.4 工作场景中 AI 威胁的影响

过往关于工作场景中 AI（失业）效应的研究取得丰富结论，然而对于“AI 失业”和个人工作之间关系的探索，还存在较大的研究空间。过往的研究或在宏观层面探讨 AI 失业可能造成的社会变化，如扩大收入差距、形成行业垄断，产生一种“掌控人工智能的少数人对不掌控人工智能的大多数人”的支配关系（Acemoglu & Autor, 2011; Lankisch et al., 2017）。或在个体层面探讨运用 AI 对职业决策的影响，例如与管理层的关系影响着员工的投入与工作表现，而在组织管理中引入 AI 会造成员工的反感并破坏原本的人际关系基础，导致员工离职意愿上升（Chughtai et al., 2015; Lee, 2018）。或关注 AI 渗透背景下人们对“失业”的态度，如相比于原先的工作岗位被人类取代，人们更能接受被 AI 取代工作的结果（Granulo et al., 2019）。或聚焦于 AI 失业引起的情绪反应，如失业焦虑、对再就业的恐惧、自我效能和幸福感的变化等等（Gimpelson & Oshchepkov, 2012; Liang & Lee, 2017）。

本文探讨了“AI 失业”与对行业和工作评价的关系，并考察了不同信息框架下的“AI 失业”如何影响人们对各类工作的评价，这在一定程度上填补了过往研究的空缺。通过实验 7 我们发现，受到概率框架的影响，即便对于需要经常接触 AI 的工作，人们对此类工作也予以更积极的评价。此外，通过对比不同信息框架下人们对 AI 研发政策的支持程度，我们发现由于概率框架降低了人们对 AI 的威胁感知，进而导致人们对开展 AI 研发政策更加支持，其中存在的 AI 威胁对相关政策支持的负向预测作用，也与此前的研究结论相一致（Złotowski et al., 2017）。

AI 威胁源自 AI 技术的进步和渗透，AI 的威胁也会引发人们的一系列反应并可能对社会发展及个人生活造成切实的影响。在面对“AI 失业”这一充满不确定性的潜在状况时，人们给予的应对方式也各不相同，其中既有如我国“瞄准人工智能等前沿领域，实施具有前瞻性、战略性的国家重大科技项目；也要培育壮大人工智能等新兴数字产业，提升通信设备、核心电子元器件、关键软件等产业水平”（人民日报, 2021）；也有如欧盟制定 AI 准入规范以限制 AI 失业的规模（Sparrow, 2007）。关注“AI 失业”，并不否认 AI 技术广泛应用所带来的就业潜能，而是着眼于技术发展的可能情况，预测和规避风险，保障 AI 在社会中健康而可持续地发展。揭示“AI 失业”的信息框架效应，则能明晰大众传媒报道、信息组织方式等改变 AI 威胁的外部因素，降低人们做出次优决策的可能。

9.5 不足与展望

本研究存在的不足与未来的方向如下。首先，不同的行业与岗位之间存在复杂的关联性，特定行业的从业者可能因为担心其他行业的劳动力被 AI 替代后流入而担心，研究 7 中

并未对参与者所处的行业进行详细编码，也没有在 AI 曝光度的基础上对行业类型和职业性质做更细致的划分，虽然 AI 曝光度与 AI 替代度存在高度相关性，但仍有个别行业可能带有低曝光高替代的性质，例如物流行业（Felten et al., 2021），未来的研究可以对此做进一步的分析。其次，AI 失业作为描述对象本身带有负面色彩，虽然我们在实验中呈现的是客观的概率数字，但 AI 对人类工作的取代仍难免引发负面的联想，因此未来的研究可以选择更加中性的表述，例如以挑战或机遇的方式来表述 AI 与人力劳动的关系，对获得-损失框架和因素-概率框架的影响进行综合的检验。再次，研究表明，在模糊容忍度之外，认知闭合需求等其他人格因素可能影响信息框架效应（Kruglanski, 1989），一些人口统计学变量则会影响个体对 AI 威胁的感知，例如与年轻人、高收入者相比，老年人、低收入者和少数族裔成员在工作中感受到更多的 AI 威胁（Ghimire et al., 2020）、组织环境中的企业高层要比基层员工感受到更低的 AI 威胁（Kolbjørnsrud et al., 2017）、女性相比于男性感受到更低的 AI 威胁等等（Gallimore et al., 2019），本研究并未对上述因素做详细的分析，未来的研究可以将此类因素纳入考察。最后，未来研究可以考察信息框架与 AI 信任的关系。AI 带给人的威胁感既有被其功用性所取代的可能（Alaiad & Zhou, 2013; Paetzel et al., 2020），也会被人们对 AI 产生的信任感所淡化（Correia et al., 2016），基于因素与概率的数字信息框架是否会直接影响人们对 AI 的信任，以及是否通过影响 AI 信任进而影响 AI 威胁，均是值得探究的问题。

10 结论

本研究使用因素信息框架和概率信息框架分别表述 AI 失业，得到以下结论：（1）AI 失业的信息框架效应显著，体现在 AI 威胁（比起因素框架，概率框架降低人们的 AI 威胁感）、对 AI 研发政策的态度（比起因素框架，概率框架增强人们对 AI 研发政策的支持）和个人的就业推荐行为（比起因素框架，概率框架提升人们对高 AI 曝光度的行业的推荐度）；（2）因素-概率框架通过影响个体的主观可能性判断（因素信息框架增加 AI 失业发生可能性的判断，概率信息框架降低 AI 失业可能性发生的判断），降低受概率框架影响者的 AI 威胁感；（3）AI 失业的信息框架效应主要表现在对模糊容忍度较低的个体当中。

参考文献

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment an d earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labour Economics* (Vol. 4, pp. 1043–1171). Amsterdam: Elsevier.

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, *128*(6), 2188–2244.

Alaiad, A., & Zhou, L. (2013, August). Patient Behavioural Intention toward Adopting Healthcare Robots. *The 19th Americas Conference on Information Systems (AMCIS)*.

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression fo rmation. *Journal of Experimental Psychology*, *70*(4), 394–400.

Baron, J. (2014). Heuristics and biases. In E. Zamir & D. Teichman (Eds.), *The Oxford handbook of behavioral economics and the law*, (pp. 3–27). Oxford University Press.

Bartneck, C., Duenser, A., Moltchanova, E., Zawieska, K., 2015. Comparing the similarity of resp onses received from studies in amazons mechanical turk to studies conducted online and with direct recruitment. *PLoS One*, *10*, 1–23.

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropo morphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Intern ational *Journal of Social Robotics*, *1*(1), 71–81.

Betsch, C., Haase, N., Renkewitz, F., & Schmid, P. (2015). The narrative bias revisited: What driv es the biasing influence of narrative information on risk perceptions? *Judgment and Decision Making*, *10*(3), 241–264.

Blanas, S., Gancia, G., & Lee, S. Y. (2019). Who is afraid of machines?. *Economic Policy*, *34*(10 0), 627–690.

Brewer, M. B. (2001). The many faces of social identity: Implications for political psychology. *Po litical Psychology*, *22*, 115–125.

Brewer, N. T., Weinstein, N. D., Cuite, C. L., & Herrington, J. E. (2004). Risk perceptions and the ir relation to risk behavior. *Annals of Behavioral Medicine*, *27*, 125–130.

Brynjolfsson, E., & McAfee, A. (2017, July 18). The business of artificial intelligence. *Harvard B usiness Review*. Retrieved May 2, 2024, from https:// hbr.org/2017/07/the–business-of-artifici

al-intelligence

Buhr, K., & Dugas, M. J. (2006). Investigating the construct validity of intolerance of uncertainty and its unique relationship with worry. *Journal of Anxiety Disorders*, *20*(2), 222–236.

Campbell, C. A. (1966). The Discipline of the Cave. *Philosophical Books*, *7*(3), 10–12.

Cheng, H., Chen, W., & Li, T. (2018). Robots in China: Status, future and impact—empirical evid ence from the China employer–employee survey (CEES). *Journal of Macro-quality Research*, *6*(3), 1–21.

Chernev, A., & Gal, D. (2010). Categorization effects in value judgments: Averaging bias in evalu ating combinations of vices and virtues. *Journal of Marketing Research*, *47*(4), 738–747.

[程虹, 陈文津, 李唐. (2018). 机器人在中国:现状、未来与影响——来自中国企业-劳动力匹配调查(CEES)的经验证据. *宏观质量研究*, *6*(3), 1–21.]

Chui, M., Manyika, J., & Miremadi, M. (2016). Where machines could replace humans—and whe re they can't (yet). *McKinsey Quarterly*. Retrieved May 2, 2024, from https://www.mckinsey.de/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Where%20machines%20could%20replace%20humans%20and%20where%20they%20cant/Where-machines-could-replace-humans-and-where-they-cant-yet.pdf

Chughtai, A., Byrne, M., & Flood, B. (2015). Linking ethical leadership to employee well-being: The role of trust in supervisor. *Journal of Business Ethics*, *128*, 653–663.

Correia, F., Alves-Oliveira, P., Maia, N., Ribeiro, T., Petisca, S., Melo, F. S., & Paiva, A. (2016). Just follow the suit! Trust in human-robot interactions during card game playing. *2016 25th I EEE International Symposium on Robot and Human Interactive Communication (ROMAN)*, 507–512.

Dang, J., & Liu, L. (2021). Robots are friends as well as foes: Ambivalent attitudes toward mindfu l and mindless AI robots in the United States and China. *Computers in Human Behavior*, *115*, 106612.

Dauth, W., Findeisen, S., Suedekum, J., & Woessner, N. (2018, August 21). Adjusting to robots: Worker-level evidence (Research Paper No. 013; Institute Working Paper). *Federal Reserve Bank of Minneapolis*. Retrieved May 2, 2024, from https://doi.org/10.21034/iwp.13

Dekker, F., Salomons, A., & Waal, J. V. D. (2017). Fear of robots at work: the role of economic se lf-interest. *Socio-Economic Review*, *15*(3), 539–562.

Diener, J. & Richardson, L.E. (2007). Attitudes Toward Seat Belt Use Among Urban & Rural Teens. *University of Missouri Columbia, Institute of Public Policy*. Retrieved May 2, 2024, from https://mospace.umsystem.edu/xmlui/handle/10355/2573

Di Tella, R., & Rodrik, D. (2020). Labour Market Shocks and the Demand for Trade Protection: Evidence from Online Surveys. *The Economic Journal*, *130*(628), 1008–1030.

Dugas, M. J., Gagnon, F., Ladouceur, R., & Freeston, M. H. (1998). Generalized anxiety disorder: A preliminary test of a conceptual model. *Behaviour Research and Therapy*, *36*(2), 215–226.

Dugas, M. J., Gosselin, P., & Ladouceur, R. (2001). Intolerance of uncertainty and worry: Investigating specificity in a nonclinical sample. *Cognitive Therapy and Research*, *25*, 551–558.

Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich.

Fan, C., Deng, Y. (2022). Workers' fear of unemployment under the background of "machines replacing humans": Evidence from matched "employer-employee" survey data for Guangdong's manufacturing industry. *Sociological Review of China*, *10*(02), 67–87.

[范长煜, 邓韵雪. (2022). “机器换人”背景下工人失业担忧研究——基于广东省制造业“企业-职工”匹配调查数据. *社会学评论*, *10*(02), 67–87.]

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, *42*(12), 2195–2217.

Frenkel-Brunswik, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *Journal of Personality*, *18*(1), 108–143.

Frey, C., & Rahbari, E. (2016). *Do labor-saving technologies spell the death of jobs in the developing world*. Paper prepared for the 2016 Brookings Blum Roundtable, Aspen, Colorado.

Furnham, A., & Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current Psychology*, *14*(3), 179–199.

Gallimore, D., Lyons, J. B., Vo, T., Mahoney, S., & Wynne, K. T. (2019). Trusting Robocop: Gender-Based Effects on Trust of an Autonomous Robot. *Frontiers in Psychology*, *10*(482), 1–9.

Ghimire, R., Skinner, J., & Carnathan, M. (2020). Who perceived automation as a threat to their jo

bs in metro Atlanta: Results from the 2019 Metro Atlanta Speaks survey. *Technology in Society*, *63*, 101368.

Gimpelson, V., & Oshchepkov, A. (2012). Does more unemployment cause more fear of unemployment?. *IZA Journal of Labor & Development*, *1*(6), 1–26.

Gnambs, T., & Appel, M. (2019). Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. *Computers in Human Behavior*, *93*, 53–61.

Granulo, A., Fuchs, C., & Puntoni, S. (2019). Psychological reactions to human versus robotic job replacement. *Nature Human Behaviour*, *3*(10), 1062–1069.

Gray, K., Yam, K. C., Zhen'An, A. E., Wilbanks, D., & Waytz, A. (2023). The psychology of robots and artificial intelligence. In D. Gilbert (Ed.), *The Handbook of Social Psychology*. Situational Press (in press).

Grenier, S., Barrette, A. M., & Ladouceur, R. (2005). Intolerance of uncertainty and intolerance of ambiguity: Similarities and differences. *Personality and Individual Differences*, *39*(3), 593–600.

Harari, Y. N. (2017). Reboot for the AI revolution. *Nature*, *550*, 324–327.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression‐based approach*. New York: Guilford Press.

Hoch, S. J., & Ha, Y. W. (1986). Consumer learning: Advertising and the ambiguity of product experience. *Journal of Consumer Research*, *13*(2), 221–233.

Hsu, A. (2022, September 8). California dockworkers are worried about losing their good-paying jobs to robots. *NPR*. Retrieved May 2, 2024, from https://www.npr.org/2022/09/08/1121690513/california-dockworkers-are-worried-about-losing-their-good-paying-jobs-to-robots

Huang, H. L., Cheng, L. K., Sun, P. C., & Chou, S. J. (2021). The effects of perceived identity threat and realistic threat on the negative attitudes and usage intentions toward hotel service robots: The moderating effect of the robot's anthropomorphism. *International Journal of Social Robotics*, *13*, 1599–1611.

Innocenti, S., & Golin, M. (2022). Human capital investment and perceived automation risks: Evidence from 16 countries. *Journal of Economic Behavior & Organization*, *195*, 27–41.

Iza World of Labor. (2019, August 29). French supermarket workers protest against automated checkout stations. *Iza World of Labor*. Retrieved May 2, 2024, from https://wol.iza.org/news/fr

ench-supermarket-workers-protest-against-automated-checkout-stations

Jackson, J. W. (1993). Realistic group conflict theory: A review and evaluation of the theoretical and empirical literature. *Psychological Record, 43,* 395– 413.

Jørgensen, C., & Navrbjerg, S. E. (2001). The involvment of employees and collective bargaining in company restructuring. *European Industrial Relations Observatory Online*. Retrieved May 2, 2024, from https://www.eurofound.europa.eu/en/industrial-restructuring

Kahn, B. (2023, March 29). Elon Musk and Apple cofounder Steve Wozniak among over 1,100 who sign open letter calling for 6-month ban on creating powerful A.I. *Fortune*. Retrieved May 2, 2024, from https://fortune.com/2023/03/29/elon-musk-apple-steve-wozniak-over-1100-sign-open-letter-6-month-ban-creating-powerful-ai/

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291.

Kareklas, I., Carlson, J. R., & Muehling, D. D. (2012). The role of regulatory focus and self-view in "green" advertising message framing. *Journal of Advertising*, *41*(4), 25–39.

Karmarkar, U. R., & Kupor, D. (2023). The unlikelihood effect: When knowing more creates the perception of less. *Journal of Experimental Psychology: General, 152*(3), 906–920.

Katsaros, K. K., & Nicolaidis, C. S. (2012). Personal traits, emotions, and attitudes in the workplace: Their effect on managers' tolerance of ambiguity. *The Psychologist-Manager Journal*, *15*(1), 37–55.

Keenan, A. (1978). Selection interview performance and intolerance of ambiguity. *Psychological Reports*, *42*(2), 353–354.

Kim, T., & Song, H. (2023). Communicating the limitations of AI: the effect of message framing and ownership on trust in artificial intelligence. *International Journal of Human—Computer Interaction*, *39*(4), 790–800.

Koch, M., Manuylov, I., & Smolka, M. (2021). Robots and Firms. *The Economic Journal*, *131*(638), 2553–2584.

Kolbjørnsrud, V., Amico, R., & Thomas, R. J. (2017). Partnering with AI: How organizations can win over skeptical managers. *Strategy & Leadership*, *45*(1), 37–43.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York, NY: Plenum Press.

Lankisch, C., Prettner, K., & Prskawetz, A. (2017). Robots and the skill premium: An automation-based explanation of wage inequality (No. 29-2017). *Hohenheim Discussion Papers in Business, Economics and Social Sciences*. Retrieved May 2, 2024, from https://www.econstor.eu/handle/10419/169370?locale=en

Lauriola, M., Russo, P. M., Lucidi, F., Violani, C., & Levin, I. P. (2005). The role of personality in positively and negatively framed risky health decisions. *Personality and Individual Differences*, *38*(1), 45–59.

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 2053951718756684.

Leo, X., & Huh, Y. E. (2020). Who gets the blame for service failures? Attribution of responsibility toward robot versus human service providers and service firms. *Computers in Human Behavior*, *113*(4), 106520.

Levin, I. P., Gaeth, G. J., Schreiber, J., & Lauriola, M. (2002). A new look at framing effects: Distribution of effect sizes, individual differences, and independence of types of effects. *Organizational Behavior and Human Decision Processes*, *88*(1), 411–429.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*, 149–188.

Li, P. (2019). Technological development and employment challenges in the era of Artificial Intelligence: A research based on the fear of unemployment risk. *Think Tank: Theory & Practice*, *4*(6), 43–51.

[李佩. (2019). 人工智能时代的技术发展与就业挑战：基于失业风险恐惧的探索. *智库理论与实践*, *4*(6), 43–51.]

Liang, Y., & Lee, S. A. (2017). Fear of autonomous robots and artificial intelligence: Evidence from national representative data with probability sampling. *International Journal of Social Robotics*, *9*, 379–384.

Lin, J., Panganiban, A. R., Matthews, G., Gibbins, K., Ankeney, E., See, C., ... & Long, M. (2022). Trust in the danger zone: Individual differences in confidence in robot threat assessments. *Frontiers in Psychology*, *13*, 601523.

Lingmont, D. N., & Alexiou, A. (2020). The contingent effect of job automating technology aware

ness on perceived job insecurity: Exploring the moderating role of organizational culture. *Technological Forecasting and Social Change*, *161*, 120302.

Lynch, J., Jr. (1979). Why additive utility models fail as descriptions of choice behavior. *Journal of Experimental Social Psychology*, *15*(4), 397–417.

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, *90*, 46–60.

McClure, P. K. (2018). “You’re fired,” says the robot: The rise of automation in the workplace, technophobes, and fears of unemployment. *Social Science Computer Review*, *36*(2), 139–156.

McElroy, T., & Seta, J. J. (2003). Framing effects: An analytic-holistic perspective. *Journal of Experimental Social Psychology*, *39*, 610–617.

McFarland, M. (2014, October 24). Elon Musk: ‘With artificial intelligence we are summoning the demon.’ *The Washington Post*. Retrieved May 2, 2024, from https://www.washingtonpost.com/news/innovations/wp/2014/10/24/elon-musk-with-artificial-intelligence-we-are-summoning-the-demon/

McLain, D. L. (1993). The MSTAT-I: A new measure of an individual's tolerance for ambiguity . *Educational and psychological measurement*, *53*(1): 183–189.

McLain, D. L. (2009). Evidence of the properties of an ambiguity tolerance measure: the multiple stimulus types ambiguity tolerance scale-ii (mstat-ii). *Psychological Reports*, *105*(3), 975–988.

Meadows, J., Catalan, J., & Gazzard, B. (1993). “I plan to have the HIV test”—Predictors of testing intention in women attending a London antenatal clinic. *AIDS Care*, *5*(2), 141–148.

Metz, C. (2022, April 6). Meet DALL-E, the A.I. that draws anything at your command. *The New York Times*. Retrieved May 2, 2024, from https://www.nytimes.com/2022/04/06/technology/openai-images-dall-e.html

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, *14*(10), 435–440.

Morikawa, M. (2017). Who are afraid of losing their jobs to artificial intelligence and robots? Evidence from a survey (Working Paper No. 71; GLO Discussion Paper). *Global Labor Organization*. Retrieved May 2, 2024, from https://www.econstor.eu/handle/10419/158005

Nemo, L. (2018, May 30). Las Vegas food service workers are going on strike so they don’t lose t

heir jobs to robots. *Futurism*. Retrieved May 2, 2024, from https://futurism.com/las-vegas-food-service-workers-strike-automation

Paetzel, M., Perugia, G., & Castellano, G. (2020). The persistence of first impressions: 15th Annual ACM/IEEE International Conference on Human Robot Interaction, HRI 2020. *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 73–82.

Padalino, S., & Vivarelli, M. (1997). The employment intensity of economic growth in the G-7 countries. *International Labor Review, 136*(2), 191–213.

Redelmeier, D. A., Koehler, D. J., Liberman, V., & Tversky, A. (1995). Probability judgement in medicine: Discounting unspecified possibilities. *Medical Decision Making*, *15*(3), 227–230.

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *10*(4), 336–353.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, *17*(5), 407–413.

Roose, K. (2022, December 5). The brilliance and weirdness of ChatGPT. *The New York Times*. Retrieved May 2, 2024, from https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html

Rotman, D. (2013). How technology is destroying jobs. *Technology Review*, *16*(4), 28–35.

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review*, *104*(2), 406–415.

Rusting, C. L., & Larsen, R. J. (1998). Personality and cognitive processing of affective information. *Personality and Social Psychology Bulletin*, *24*(2), 200–213.

Santoro, E., & Monin, B. (2023). The AI effect: People rate distinctively human attributes as more essential to being human after learning about artificial intelligence advances. *Journal of Experimental Social Psychology, 107*, 104464.

Samuel, S. (2020, January 13). Robot priests can bless you, advise you, and even perform your funeral. *Vox*. Retrieved May 2, 2024, from https://www.vox.com/future-perfect/2019/9/9/20851753/ai-religion-robotpriest-mindar-buddhism-christianity

Savage, N. (2020). How AI is improving cancer diagnostics. *Nature, 579*(7800), S14–S16.

Schlogl, L., & Sumner, A. (2018, July 02). The rise of the robot reserve army: Automation and the future of economic development, work, and wages in developing countries. *Center for Globa*

l Development Working Paper 487. Retrieved May 2, 2024, from https://www.cgdev.org/publication/rise-robot-reserve-army-automation-and-future-economic-development-work-and-wages

Shoss, M. K., & Ciarlante, K. (2022). Are robots/AI viewed as more of a workforce threat in unequal societies? Evidence from the eurobarometer survey. *Technology, Mind, and Behavior*, *3*(2), 1–13.

Sparrow, R. (2007). Killer robots. *Journal of Applied Philosophy*, *24*(1), 62–77.

Stanovich, K. E., & West, R. F. (1998). Individual differences in framing and conjunction effects. *Thinking and Reasoning*, *4*(4), 289–317.

Stephan, W. G., Ybarra, O., & Rios, K. (2015). Intergroup threat theory. In *Handbook of Prejudice, Stereotyping, and Discrimination* (pp. 255 – 278). Psychology Press.

Strogatz, S. (2018, December 26). One giant step for a chess-playing machine. *The New York Times*. Retrieved May 2, 2024, from https://wwidentityw.nytimes.com/2018/12/26/science/chess-artificial-intelligence.html

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of Intergroup Relations* (pp. 7–24). Chicago, IL: Nelson-Hall.

Tang, P. M., Koopman, J., Yam, K. C., De Cremer, D., Zhang, J. H., & Reynders, P. (2022). The self-regulatory consequences of dependence on intelligent machines at work: Evidence from field and experimental studies. *Human Resource Management*, 1–24.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

Tversky, A., & Koehler, D. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*(4), 547–567.

Verma, P. (2022, December 1). Meta's new AI is skilled at a ruthless, power-seeking game. *The Washington Post*. Retrieved May 2, 2024, from https://www.washingtonpost.com/technology/2022/12/01/meta-diplomacy-ai-cicero/

Vincent, J. (2022, September 2). An AI-generated artwork's state fair victory fuels arguments over 'what art is.' *The Verge*. Retrieved May 2, 2024, from https://www.theverge.com/2022/9/1/23332684/ai-generatedartwork-wins-state-fair-competition-colorado

Waardenburg, L., Huysman, M., & Sergeeva, A. V. (2022). In the land of the blind, the one-eyed man is king: Knowledge brokerage in the age of learning algorithms. *Organization Science*, *33*(1), 59–82.

Weaver, K., & Garcia, S. M. (2018). The adding-and-averaging effect in bundles of information: Preference reversals across joint and separate evaluation. *Journal of Experimental Psychology: Applied*, *24*(3), 296–305.

Webster, D. M. & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*, 1049–1062.

Wen, G., Xu, F., Yu, H., Huang, B., & Wang, L. (2011). The psychological mechanism and influential factors of attribute framing effect. *Advances in Psychological Science*, *19*(12), 1822–1833.

[文桂婵, 徐富明, 于会会, 黄宝珍, 王岚. (2011). 特征框架效应的心理机制与影响因素. *心理科学进展*, *19*(12), 1822–1833.]

Wike, R., & Stokes, B. (2018). In advanced and emerging economies alike, worries about job automation. *Pew Research Center, Global Attitudes & Trends*. Retrieved May 2, 2024, from https://www.pewglobal.org/wp-content/uploads/sites/2/2018/09/Pew-Research-Center_In-Advanced-and-Emerging-Economies-Alike-Worries-about-Job-Automation_2018-09-13.pdf

Wong, R. S. (2020). An alternative explanation for attribute framing and spillover effects in multidimensional supplier evaluation and supplier termination: Focusing on asymmetries in attention. *Decision Sciences*, *52*(2), 262–282.

Xu, H., & Tracey, T. J. (2014). The role of ambiguity tolerance in career decision making. *Journal of Vocational Behavior*, *85*(1), 18–26.

Xu, H., & Tracey, T. J. (2015). Ambiguity tolerance with career indecision: An examination of the mediation effect of career decision-making self-efficacy. *Journal of Career Assessment*, *23*(4), 519–532.

Xu, L., Yu, F. (2020). Factors that influence robot acceptance. *Chinese Science Bulletin*, *65*(6), 496–510.

[许丽颖, 喻丰. (2020). 机器人接受度的影响因素. *科学通报*, *65*(6), 496–510.]

Xu, L., Yu, F., Peng, K. (2021). Algorithmic discrimination causes less desire for moral punishment than human discrimination. *Acta Psychologica Sinica*, *54*(9), 1076–1092.

[许丽颖, 喻丰, 彭凯平. (2021). 算法歧视比人类歧视引起更少道德惩罚欲. *心理学报*, *54*(9), 1076–1092.]

Yam, K. C., Tang, P. M., Jackson, J. C., Su, R., & Gray, K. (2023). The rise of robots increases jo

b insecurity and maladaptive workplace behaviors: Multimethod evidence. *Journal of Applied Psychology*, *108*(5), 850–870.

Yogeeswaran, K., Złotowski, J., Livingstone, M., Bartneck, C., Sumioka, H., & Ishiguro, H. (2016). The interactive effects of robot anthropomorphism and robot ability on perceived threat and support for robotics research. *Journal of Human-Robot Interaction*, *5*(2), 29–47.

Złotowski, J., Yogeeswaran, K., & Bartneck, C. (2017). Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies*, *100*, 48–54.

中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要. *人民日报*, 01.

The Information Framing Effect of "AI Unemployment"

Abstract

The advancement of artificial intelligence (AI) technology significantly contributes to enhancing productivity; however, concerns regarding potential technological unemployment have garnered considerable attention. The uncertainty surrounding the occurrence, timing, and scale of AI-induced unemployment impedes definitive conclusions. This uncertainty may also lead the public to be influenced by encountered information concerning AI-induced unemployment. Media coverage on AI-induced unemployment often presents extensive information regarding affected industries, occupations, and probability scales, establishing two numerical information frameworks: one emphasizing factors influencing unemployment distribution across industries and another emphasizing the probability of unemployment occurrence. Comparatively, the probability framework, as opposed to the factor framework, allows individuals to formulate judgments indicating a reduced likelihood of AI-induced unemployment, thereby mitigating the perceived threat of AI, especially among individuals with high ambiguity tolerance.

Building upon the foundational assumption that the probability framework alleviates AI threat perception, this study, comprising seven recursive experiments, investigates the mediating role of judgments on AI-induced unemployment likelihood and the moderating role of individual ambiguity tolerance. Experiment 1 juxtaposes AI threat perception elicited by general AI-induced unemployment descriptions, factor frameworks, and probability frameworks. Experiment 2 validates the mediating role of likelihood judgments. Experiments 3 and 4 respectively eliminate potential influences of probability values and unemployment scale. Experiment 5 explores ambiguity tolerance's moderating effect. Experiments 6 and 7 examine subsequent AI threat effects, including support for AI development policies and willingness to recommend various occupations.

The primary findings are as follows. Firstly, introducing AI-induced unemployment through a probability framework effectively diminishes AI threat perception (Experiments 1-7). Secondly, this effect is mediated by perceived likelihood, whereby the probability framework prompts individuals to form judgments indicating decreased AI-induced unemployment likelihood, thus reducing AI threat (Experiments 2-5). Thirdly, the information framework effect is moderated by ambiguity tolerance, primarily manifesting among individuals tolerant of ambiguous information (Experiment 5). Fourthly, individuals influenced by the probability framework demonstrate increased support for policies related to AI development, with AI threat playing a mediating role (Experiment 6). Lastly, individuals influenced by the probability framework exhibit a heightened willingness to recommend jobs involving frequent AI interaction (Experiment 7).

This study extends prior research by elucidating how external factors such as information frames contribute to variations in AI threat perception. Unlike the extensively studied valence information frame, numerical information frames impact AI threat perception by altering individuals' likelihood judgments. Our findings shed light on the effects of the numerical information framework on AI-induced unemployment threat perception, policy support, and job recommendation willingness.

Keywords technological unemployment; information processing; probability judgments; perceived AI

threat; ambiguity tolerance